

Random matrices, log-gases and Hölder regularity

László Erdős*

May 15, 2014

Abstract. The Wigner-Dyson-Gaudin-Mehta conjecture asserts that the local eigenvalue statistics of large real and complex Hermitian matrices with independent, identically distributed entries are universal in a sense that they depend only on the symmetry class of the matrix and otherwise are independent of the details of the distribution. We present the recent solution to this half-century old conjecture. We explain how stochastic tools, such as the Dyson Brownian motion, and PDE ideas, such as De Giorgi-Nash-Moser regularity theory, were combined in the solution. We also show related results for log-gases that represent a universal model for strongly correlated systems. Finally, in the spirit of Wigner's original vision, we discuss the extensions of these universality results to more realistic physical systems such as random band matrices.

Mathematics Subject Classification (2010). Primary 15B52; Secondary 82B44.

Keywords. De Giorgi-Nash-Moser parabolic regularity, Wigner-Dyson-Gaudin-Mehta universality, Dyson Brownian motion

1. Introduction

Large complex systems with many degrees of freedom often exhibit remarkably simple universal patterns. The Gauss law describes the fluctuations of large sums of independent or weakly dependent random variables irrespective of their distribution. The Poisson point process is the universal model for many independent events in space or time. Both laws are ubiquitous in Nature thanks to their large domain of attraction but they cannot accurately model strong correlations. Can one find a universality for correlated systems?

Since correlations appear in many forms, this seems an impossible task. Nevertheless this is exactly what E. Wigner has accomplished when he discovered a universal pattern in the spectrum of heavy nuclei. Spectral measurement data for various nuclei clearly show that the density of energy levels depends on the actual nucleus. But Wigner asked a different question: he looked at the energy *gaps*, i.e.

*Partially supported by SFB-TR 12 Grant of the German Research Council and by ERC Advanced Grant, RANMAT 338804.

the *difference* of consecutive energy levels. He discovered that their statistics, after rescaling with the local density, showed a very similar pattern for different nuclei.

Wigner's revolutionary insight was that this coincidence does not stem from some particular property of the specific physical system but it has a profound mathematical origin. General quantum mechanics postulates that energy levels are eigenvalues of a certain hermitian matrix (or operator) $H = (h_{ij})$, the *Hamiltonian* of the system. The matrix elements h_{ij} represent quantum transition rates between two states labelled by i and j . While h_{ij} 's are specific to the system, the gap statistics largely depend only on the basic symmetry class of H , as long as h_{ij} 's are chosen somewhat generically.

To illustrate this mechanism, consider a 2×2 hermitian matrix

$$H = \begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix}, \quad a, d \in \mathbb{R}, \quad b \in \mathbb{C}.$$

The difference (or gap) of the two eigenvalues is $\lambda_2 - \lambda_1 = [(a - d)^2 + 4|b|^2]^{1/2}$. If the matrix elements are drawn independently from some continuous distribution, then the probability that the gap is very small;

$$\mathbb{P}(|\lambda_2 - \lambda_1| \leq \varepsilon), \quad \varepsilon \ll 1,$$

is of order ε^2 for real symmetric matrices ($b \in \mathbb{R}$) and it is of order ε^3 for complex hermitian matrices ($b \in \mathbb{C}$). The exponent of ε is thus determined by the symmetry class of H .

Very surprisingly, for large $N \times N$ matrices the *entire distribution* of the gap becomes universal as $N \rightarrow \infty$ and not only its asymptotics in the $\varepsilon \ll 1$ regime. Based upon a more precise calculation with Gaussian matrix elements, Wigner predicted that this universal law is given by a simple formula (called the *Wigner surmise*). For the real symmetric case it is

$$\mathbb{P}(\tilde{\lambda}_j - \tilde{\lambda}_{j-1} = s + ds) \approx \frac{\pi s}{2} \exp\left(-\frac{\pi}{4}s^2\right)ds, \quad (1)$$

where $\tilde{\lambda}_j = \varrho \lambda_j$ denote the eigenvalues λ_j rescaled by the density of eigenvalues ϱ near λ_j . This law is characteristically different from the gap distribution of the Poisson point process which is the exponential distribution, $e^{-s}ds$. The prefactor s in (1) indicates a *level repulsion* for the point process $\tilde{\lambda}_j$, in particular the eigenvalues are strongly correlated (eigenvalues are often called (energy) levels in random matrix theory). Similar formulas hold for the joint statistics of several consecutive gaps.

Comparing measurement data from various experiments, Wigner concluded that the energy gap distribution of complicated quantum systems is essentially universal; it depends only on the basic symmetries of model (such as time-reversal invariance). This thesis has never been rigorously proved for any realistic physical system but experimental data and extensive numerics leave no doubt on its correctness (see [50] for an overview).

Once universality is expected, explicit formulas for the statistics can be computed from the most convenient model within the universality class. The simplest

representatives of these universality classes are $N \times N$ random matrices with independent (up to symmetry), identically distributed Gaussian entries. These are called the *Gaussian orthogonal ensemble (GOE)* and the *Gaussian unitary ensemble (GUE)* in case of real symmetric and complex Hermitian matrices, respectively. Wigner's bold vision was to neglect all details of the actual Hamiltonian operator and replace it with a large random Gaussian matrix of the same symmetry class. As far as the gap statistics are concerned, this simple-minded model very accurately reproduced the behavior of large complex quantum systems!

Since Wigner's discovery random matrix statistics are found everywhere in physics and beyond, wherever nontrivial correlations prevail. Random matrix theory (RMT) is present in chaotic quantum systems in physics, in principal component analysis in statistics, in communication theory and even in number theory. In particular, the zeros of the Riemann zeta function on the critical line are expected to follow RMT statistics due to a spectacular result of Montgomery [54].

In retrospect, Wigner's idea should have received even more attention. For centuries, the primary territory of probability theory was to model uncorrelated or weakly correlated systems. The surprising ubiquity of random matrix statistics is a strong evidence that it plays a similar fundamental role for correlated systems as Gaussian distribution and Poisson point process play for uncorrelated systems. RMT seems to provide essentially the only universal and generally computable pattern for complicated correlated systems.

A few years after Wigner's seminal paper [75], Gaudin [41] discovered another remarkable property of this new point process: the correlation functions have an exact determinantal structure, at least if the distributions of the matrix elements are Gaussian. The algebraic identities within the determinantal form opened up the route to obtain explicit formulas for local correlation functions. For example, in the complex Hermitian case (GUE) the n -point correlation function $p^{(n)}$ of the rescaled eigenvalues $\tilde{\lambda}_i$ in the bulk is given by the determinant of the celebrated sine-kernel:

$$p^{(n)}(\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_n) = \det[K(\tilde{\lambda}_i, \tilde{\lambda}_j)]_{i,j=1}^n, \quad K(x, y) := \frac{\sin \pi(x - y)}{\pi(x - y)}. \quad (2)$$

(The same determinantal expression with a different but closely related kernel function K holds for the real symmetric case.) As a consequence, the gap distribution is given by a Fredholm determinant involving Hermite polynomials. In fact, Hermite polynomials were first introduced in the context of random matrices by Mehta and Gaudin [52] earlier. Dyson and Mehta [51, 19, 21] have later extended this exact calculation to correlation functions and to other symmetry classes. When compared with the exact formula, the Wigner surmise (1), based upon a simple 2×2 matrix model, turned out to be quite accurate. While the determinantal structure is present only in Gaussian Wigner matrices, the paradigm of spectral universality predicts that the formulas for the local eigenvalue statistics obtained in the Gaussian case hold for general distributions as well.

2. Random matrix ensembles and log-gases

We consider $N \times N$ hermitian matrices H with matrix elements having mean zero and variance $1/N$, i.e.

$$\mathbb{E} h_{ij} = 0, \quad \mathbb{E} |h_{ij}|^2 = \frac{1}{N} \quad i, j = 1, 2, \dots, N. \quad (3)$$

The matrix elements h_{ij} are real or complex independent random variables subject to the symmetry constraint $h_{ij} = \bar{h}_{ji}$. These ensembles of random matrices are called (*standard*) *Wigner matrices*. The normalization (3) is introduced for definiteness.

An important special case of Wigner matrices is the Gaussian case (GOE or GUE), when h_{ij} 's have Gaussian distribution. In this case the matrix ensemble can also be given by the probability law

$$P(H) dH = Z^{-1} e^{-\frac{\beta}{4} N \operatorname{Tr} H^2} dH, \quad (4)$$

where $dH = \prod_{i < j} dh_{ij} d\bar{h}_{ij} \prod_i dh_{ii}$ is the standard Lebesgue measure on real symmetric or complex hermitian $N \times N$ matrices and $Z = Z_N$ is the normalization. The parameter β is chosen to be $\beta = 1$ for GOE and $\beta = 2$ for GUE to ensure the normalization (3).

The representation (4) shows that the Gaussian ensembles enjoy an invariance property; the distribution $P(H)$ is invariant under a base transformation, $H \rightarrow U H U^*$, where U is orthogonal (in case of GOE) or unitary (in case of GUE). In fact, invariance property is not restricted to the Gaussian case; one may directly generalize (4) to

$$P(H) dH = Z^{-1} e^{-\frac{\beta}{2} N \operatorname{Tr} V(H)} dH, \quad (5)$$

where $V : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary function with sufficient growth at infinity to ensure the normalizability of the measure. The ensembles of the form (5) are called *invariant matrix ensembles*.

Wigner ensembles and invariant ensembles represent two natural but quite different ways to equip the space of $N \times N$ matrices with a probability measure. These two families are essentially disjoint; only the Gaussian ensembles belong to their intersection.

Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_N)$ denote the eigenvalues of H in increasing order. Since eigenvalues are complicated functions of the matrix elements, there is no explicit formula to express the probability distribution of $\boldsymbol{\lambda}$ induced by a general Wigner ensemble. However, quite remarkably, for invariant ensembles (5) the joint probability density of the eigenvalues is explicitly given by

$$\mu_{\beta, V}^{(N)}(\boldsymbol{\lambda}) = C \prod_{1 \leq i < j \leq N} (\lambda_j - \lambda_i)^\beta \prod_{j=1}^N e^{-\frac{\beta}{2} N V(\lambda_j)} \quad (6)$$

with a normalization constant C . This formula may directly be obtained from (5) by diagonalizing $H = U \Lambda U^*$ and integrating out the matrix of eigenvectors $U \in O(N)$ or $U \in U(N)$ with respect to the Haar measure.

From statistical physics point of view, we may consider the distribution (6) as a Gibbs measure for a gas of N point particles on \mathbb{R} . We may write

$$\mu_{\beta,V}^{(N)}(\boldsymbol{\lambda}) = C e^{-\beta N \mathcal{H}(\boldsymbol{\lambda})}, \quad \mathcal{H}(\boldsymbol{\lambda}) := \sum_{k=1}^N \frac{1}{2} V(\lambda_k) - \frac{1}{N} \sum_{1 \leq i < j \leq N} \log(\lambda_j - \lambda_i), \quad (7)$$

where, according to the Gibbs formalism, $\mathcal{H}(\boldsymbol{\lambda})$ is the Hamiltonian (energy function) of the gas and the parameter β plays the role of the inverse temperature. The Vandermonde determinant in (6) translates into a logarithmic pair interaction between the particles. We may completely ignore the original random matrix ensemble behind (6) and consider (7) more generally for any parameter $\beta > 0$, not only for the specific values $\beta = 1, 2$. The Gibbs measure (7) is often called *β -log-gas* or *β -ensemble*.

Eigenvalue distributions of Wigner ensembles and β -log-gases are quite different mathematical entities despite their connection via (6) in the special Gaussian case, $V(\lambda) = \frac{1}{2}\lambda^2$ and $\beta = 1, 2$. Wigner ensembles are parametrized by the value $\beta = 1, 2$ and by the distribution of the single matrix elements, while log-gases are parametrized by β and the potential function V . The central thesis of universality asserts that the gap statistics of both families of ensembles depend only on the parameter β and are otherwise independent of any other details of the models.

For Wigner matrices this thesis is generally referred to as the *universality conjecture of random matrices* and we will call it the *Wigner-Dyson-Gaudin-Mehta conjecture*. It was first formulated in Mehta's treatise on random matrices [50] in 1967 and has remained a key question in the subject ever since. In this article we review the recent progress that has led to the proof of this conjecture and the analogous conjecture for log gases. For more details, the reader is referred to the lecture notes [22].

3. Random band matrices and Anderson model

As mentioned in the introduction, Wigner's vision extends the thesis of universality far beyond the models we just introduced. We now present an extension that was an important source of motivation in the development of the subject.

Viewed as a quantum mechanical Hamilton operator, a Wigner matrix H represents a *mean-field system*; the quantum transition rates h_{ij} between any two quantum states, labelled by i and j , are comparable in size. The quantum states of more realistic physical models have a spatial structure and typically quantum transition occurs between nearby states only.

The spatial structure is essential to understand the *metal-insulator transition* which is the fundamental phase transition of disordered quantum systems modelled by a random Hamilton operator H . According to the physical theory, in the *metallic phase* the eigenfunctions are delocalized, the quantum time evolution e^{itH} is diffusive and the local eigenvalue statistics coincide with the ones from the GUE/GOE random matrix theory (2). The *localization length*, which is the characteristic lengthscale of the physically relevant quantities (such as eigenfunctions or

propagators), is practically infinite. In contrast, in the *insulator phase*, the eigenfunctions are localized with a localization length ℓ independent of the system size, the time evolution remains bounded for all times and the local eigenvalue statistics are Poisson. In the mathematics literature these two phases are usually called *delocalized* and *localized* regimes, respectively, and they are primarily characterized by the spectral type (absolutely continuous vs. pure point) of the corresponding infinite volume operator.

The basic model for the metal-insulator transition is the celebrated Anderson model in solid state physics [3]. The Anderson Hamiltonian is given by $-\Delta + V(x)$ on the Hilbert space $\ell^2(\mathbb{Z}^d)$, where Δ is the lattice Laplacian and $V(x)$ is a real valued random potential field such that $\{V(x) : x \in \mathbb{Z}^d\}$ are independent and identically distributed centered random variables with variance $\sigma^2 := \mathbb{E}|V(x)|^2$. The Anderson model has been extensively studied mathematically. In nutshell, the high disorder regime is relatively well understood since the seminal work of Fröhlich and Spencer [39] for localization (an alternative proof is given by Aizenman and Molchanov [1]), complemented by the work of Minami [53] proving the local Poissonian spectral statistics. In contrast, in the low disorder regime, starting from three spatial dimension and away from the spectral edges, the eigenfunctions are conjectured to be delocalized but no rigorous proof exists (*extended states conjecture*).

Random band matrices are another popular model for the metal-insulator transition [68]. For definiteness, let the state space be a finite box $\Lambda := [1, L]^d \subset \mathbb{Z}^d$ of the d -dimensional integer lattice equipped with periodic boundary condition. We consider hermitian matrices $H = (h_{ij})_{i,j \in \Lambda}$ whose rows and columns are labelled by the elements of Λ and whose matrix elements are independent. Given a parameter $W \leq L/2$, called the *band width*, we assume that the matrix elements h_{ij} vanish beyond a distance $|i - j| \geq W$, i.e. we replace (3) with the condition

$$\mathbb{E} h_{ij} = 0, \quad \forall i, j \in \Lambda; \quad \text{and} \quad h_{ij} = 0 \quad \text{for} \quad |i - j| \geq W. \quad (8)$$

($|\cdot|$ denotes the periodic distance on Λ). These are called *random band matrices*. We often assume a translation invariant profile for the variances, i.e. that

$$\sigma_{ij}^2 := \mathbb{E}|h_{ij}|^2 = \frac{1}{W^d} f\left(\frac{|i - j|}{W}\right) \quad (9)$$

with some compactly supported function $f \geq 0$ on \mathbb{R}^d with $\int_{\mathbb{R}^d} f = 1$. Notice that the normalization is chosen such that

$$\sum_{j \in \Lambda} \sigma_{ij}^2 = 1, \quad \forall i \in \Lambda. \quad (10)$$

If the band width is maximal, $W = L/2$, and f is constant on $[-\frac{1}{2}, \frac{1}{2}]^d$, then we recover the Wigner matrices (3). Wigner matrices are always in the delocalized regime as it was shown that all eigenfunctions are extended with very high probability [31]. The other extreme is when W remains bounded even as the matrix size $|\Lambda| = L^d$ goes to infinity. This system behaves very similarly to the Anderson

model. In particular, in $d = 1$ it exhibits Anderson localization even if W grows slowly with L as $W \ll L^{1/8}$ [60]. Therefore random band matrices with an intermediate band width, $1 \ll W \ll L$, serve as a model to study the metal-insulator transition. The fundamental conjecture in $d = 1$ is that the transition occurs at $W = L^{1/2}$. This conjecture is supported by supersymmetric (SUSY) functional integration techniques [40] which are intriguingly elegant but notoriously hard to justify with full mathematical rigour. Nevertheless, very recently sine-kernel local statistics (2) were proven for a Gaussian band matrix with a specifically chosen block structure [65] using SUSY approach. The details have been worked out for $W \geq L^{1-\varepsilon}$ with some small $\varepsilon > 0$. In a related problem (correlation function of the characteristic polynomial of H at two different energies) the result even holds down to the critical band width $W \geq L^{1/2+\varepsilon}$ [64], but still only for a specific block structure and Gaussian distribution.

For more general band matrices the universality of the local statistics have not yet been proven, but it was shown in $d = 1$ that the localization length is at least $W^{5/4}$, indicating band matrices with band width at least $W \gg L^{4/5}$ are in the delocalized regime [29].

4. Universality on three levels

We consider an ensemble of N (unordered) random points $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$ on the real line, either given by eigenvalues of hermitian random matrices or points of a log-gas. We always choose the normalization such that all points lie in a bounded interval, independent of N , with a very high probability. The typical spacing between the points is therefore of order $1/N$.

The statistics of λ are characterized by the n -point functions $p_N^{(n)}$. They are defined by the following relation that holds for any function O of n variables:

$$\begin{aligned} \mathbb{E} \binom{N}{n}^{-1} \sum^* O(\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_n}) \\ = \int_{\mathbb{R}^n} p_N^{(n)}(x_1, x_2, \dots, x_n) O(x_1, \dots, x_n) dx_1 \dots dx_n. \end{aligned} \quad (11)$$

Here the star indicates that the summation runs over all n -tuples of distinct integers, (i_1, i_2, \dots, i_n) with $1 \leq i_j \leq N$. The correlation function for $n = 1$ is called the *density*. Typically we fix n and consider the limit of the correlation functions $p_N^{(n)}$ as $N \rightarrow \infty$ to obtain the limiting statistics.

We may consider the limiting statistics of the points on three scales. For definiteness we illustrate these scales for Wigner matrices; similar results hold for the log-gases and for random band matrices, but the latter only under more restrictive conditions.

4.1. Macroscopic scale. The largest scale corresponds to observable functions O in (11) that are unscaled with N . For Wigner matrices (3) the limiting

density is given by the celebrated semicircle law [75]

$$\varrho_{sc}(x) := \frac{1}{2\pi} \sqrt{(4 - x^2)_+} \quad (12)$$

in the form of a weak limit:

$$\mathbb{E} \frac{1}{N} \sum_i O(\lambda_i) = \int_{\mathbb{R}} p_N^{(1)}(x) O(x) dx \rightarrow \int_{\mathbb{R}} \varrho_{sc}(x) O(x) dx, \quad \text{as } N \rightarrow \infty, \quad (13)$$

that holds for any continuous, compactly supported function O . In fact, the semicircle law also holds not only in expectation but also as a convergence in probability for the empirical density:

$$\mathbb{P} \left(\left| \frac{1}{N} \sum_i O(\lambda_i) - \int_{\mathbb{R}} \varrho_{sc}(x) O(x) dx \right| \geq \varepsilon \right) \rightarrow 0 \quad (14)$$

for any $\varepsilon > 0$ as $N \rightarrow \infty$.

These results are the simplest form of spectral universality; they assert that the eigenvalue density on macroscopic scales is independent of the specific distribution of the matrix elements. In fact, this result also holds for *generalized Wigner matrices* whose matrix elements are still centered and independent, but their distributions may vary. The semicircle law (13) holds as long as the row sums of the variances is constant, i.e.

$$\sum_j \sigma_{ij}^2 = 1, \quad \sigma_{ij} := \mathbb{E} |h_{ij}|^2, \quad (15)$$

for any i . If (15) does not hold but the variances have a macroscopic profile in a sense that

$$\sigma_{ij}^2 = S\left(\frac{i}{N}, \frac{j}{N}\right)$$

with some fixed function S on $[0, 1] \times [0, 1]$, then the limiting density still exists and can be computed from S , but it is not given by the semicircle law any more [2]. These results show that the limiting density is determined by variances of the matrix elements alone and not by their full distribution.

4.2. Mesoscopic scales. We now consider an N -dependent scaling parameter $\eta = \eta_N > 0$ and a fixed point E in the support of the limiting density, $|E| < 2$ (real numbers E in the context of location in the spectrum are often called *energy* due to the physical meaning of the spectrum). The regime $1/N \ll \eta \ll 1$ corresponds to mesoscopic scales; on these scales the fluctuation of the empirical density around the semicircle density profile is still negligible, but the effects of individual points are not yet visible.

We rescale the observable around E in a window of size η and consider

$$\mathbb{E} \frac{1}{N\eta} \sum_i O\left(\frac{\lambda_i - E}{\eta}\right) = \int_{\mathbb{R}} p_N^{(1)}(E + x\eta) O(x) dx. \quad (16)$$

If $\eta \rightarrow 0$ as $N \rightarrow \infty$, then formally (13) would indicate that the limit of (16) is $\varrho_{sc}(E) \int O(x) dx$. This is indeed correct, with some technical assumptions even in the stronger sense (14), as long as $1/N \ll \eta \ll 1$. This is called the *local semicircle law* in the bulk of the spectrum. The first result down to the optimal scale $\eta \gg 1/N$ (modulo $\log N$ factors) was given in [31] followed by several improvements and generalizations, see [27] for a summary. In particular, local semicircle law has also been extended to the spectral edge, $|E| = 2$, where the optimal scale is $\eta \gg N^{-2/3}$ reflecting the fact that the eigenvalue spacing near the edge is of order $N^{-2/3}$.

Local semicircle laws imply, among others, that the points λ_j are very close to their *classical location* denoted by γ_j and defined as the j -th quantile of the limiting density:

$$\int_{-\infty}^{\gamma_j} \varrho_{sc}(x) dx = \frac{j}{N}. \quad (17)$$

More precisely, we have for any j (including the extreme eigenvalues near the spectral edge) that

$$|\lambda_j - \gamma_j| \lesssim |\gamma_{j+1} - \gamma_j| \quad (18)$$

with a very high probability, where \lesssim indicates logarithmic factors [37]. The property (18) is called *rigidity* and it asserts that the fluctuation of the points is essentially on the scale of the local gap $|\gamma_{j+1} - \gamma_j|$. In particular, for points in the bulk spectrum, their fluctuation is only slightly larger than $1/N$.

Local semicircle law also holds for random band matrices with (9), however the local density is controlled only down to scales $\eta \gg W^{-1}$, see [27] for a summary and also [66]. The regime $\eta \ll W^{-1}$ is mathematically unexplored and there is no optimal rigidity result.

While the density on mesoscopic scales behaves exactly as on the macroscopic scale, the density-density correlation exhibits a new universality. For two random variables, X, Y , let $\langle X; Y \rangle = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$ denote their covariance. Consider two energies $E_2 \geq E_1$ and a scale η such that $N^{-1/7} \ll \eta \ll E_2 - E_1 \ll 1$. Then for Wigner matrices the covariance decays with a universal power-law [7, 25, 26]

$$\left\langle \frac{1}{N\eta} \sum_i O\left(\frac{\lambda_i - E_1}{\eta}\right); \frac{1}{N\eta} \sum_i O\left(\frac{\lambda_i - E_2}{\eta}\right) \right\rangle \sim -[N(E_2 - E_1)]^{-2} \quad (19)$$

(for Gaussian case the result extends to $\eta \gg 1/N$ [6]). Higher order moments satisfy the Wick theorem asymptotically, i.e. the local densities at different energies converge to a Gaussian variables with a non-trivial covariance structure [25, 26].

Similar result holds for band matrices with (9) in d dimensions, but the power law decay in (19) undergoes a phase transition. For $W^{-d/7} \ll \eta \ll (W/L)^2$ the asymptotics (19) holds with the mean-field exponent -2 , while for $(W/L)^2 \ll \eta \ll 1$ the power in the right hand side becomes $-2 + \frac{d}{2}$ for $d = 1, 2, 3$ and it is logarithmic for $d = 4$. In higher dimensions, $d \geq 5$, the universality breaks down. This feature is closely related to the quantum diffusion phenomenon for the unitary time evolution [23, 24]. In the physics literature these asymptotics are called the Altshuler-Shklovskii formulas and recently they have been rigorously proved [25, 26].

4.3. Microscopic scale. The most intriguing regime for universality is the microscopic scale where the scaling parameter η in the observable is chosen comparable with the typical local eigenvalue spacing. In particular, individual eigenvalues are observed. This is the regime for the gap distribution in Wigner's surmise, and the original conjecture of Mehta [50] on random matrix universality also pertains microscopic scales.

Before we formulate the precise results, we make two remarks to explain why there will be different universality theorems.

First, for the local statistics we need to distinguish the bulk spectrum where $\eta \sim 1/N$ and the edge spectrum where $\eta \sim N^{-2/3}$. Not only the scaling but also the explicit formulas are different in these two regimes. The correlation functions are asymptotically determinantal (Pfaffian) in both cases, but in the bulk they are given by the Dyson sine kernel (2) and its real symmetric counterpart, while at the edge they are given by the Airy kernel [72, 73]. In all cases the explicit formulas have been computed in the corresponding Gaussian model which is computationally the most accessible case via orthogonal polynomials. The significance of orthogonal polynomials in random matrices has first been realized by Gaudin, Mehta and Dyson [41, 52, 21]. Their approach was later generalized and combined with the Riemann-Hilbert method to yield explicit asymptotic calculations for broader classes of invariant ensembles, see [5, 14, 17, 18, 38, 49, 57, 58] for the extensive literature in the $\beta = 2$ case and [16, 47, 62] for the more complicated $\beta = 1, 4$ case. Our universality results show that the local statistics for a general Wigner matrix or invariant ensemble (or even more generally a β -log-gas) coincide with those of the corresponding Gaussian model. Therefore all explicit asymptotic calculations apart from the simplest Gaussian case become redundant.

Second, there is a subtle difference between the universality of n -point local correlation functions around a *fixed energy* E and the universality of n consecutive points $\lambda_{j+1}, \lambda_{j+2}, \dots, \lambda_{j+n}$ for some *fixed label* j . The former asks for identifying the limit

$$\begin{aligned} \mathbb{E} \frac{1}{(N\eta)^n} \sum^* O\left(\frac{\lambda_{i_1} - E}{\eta}, \frac{\lambda_{i_2} - E}{\eta}, \dots, \frac{\lambda_{i_n} - E}{\eta}\right) \\ = \int_{\mathbb{R}^n} p_N^{(n)}(E + x_1\eta, E + x_2\eta, \dots, E + x_n\eta) O(x_1, \dots, x_n) dx_1 \dots dx_n \end{aligned} \quad (20)$$

for any smooth, compactly supported observable O , i.e. identifying the weak limit of the rescaled correlation functions $p_N^{(n)}(E + x_1\eta, E + x_2\eta, \dots)$ in the variables x_1, \dots, x_n . The latter asks for the joint distribution of $\lambda_{j+1}, \lambda_{j+2}, \dots, \lambda_{j+n}$ with an appropriate rescaling.

The rigidity (18) locates the j -th eigenvalue λ_j around a fixed energy $E = \gamma_j$ but only with a precision slightly larger than $1/N$. In fact, for the Gaussian ensembles it is known [42, 56] that $\lambda_j - \gamma_j$ is Gaussian and it fluctuates on scale $\sqrt{\log N}/N$ therefore there is no direct translation between the two types of universality results. In particular, the universality of n consecutive gaps which was

originally advocated by Wigner, i.e. the limit of

$$\begin{aligned} \mathbb{E} O\left(\frac{\lambda_{j+1} - \lambda_j}{\eta}, \frac{\lambda_{j+2} - \lambda_{j+1}}{\eta}, \dots, \frac{\lambda_{j+n} - \lambda_{j+n-1}}{\eta}\right) \\ = \int_{\mathbb{R}^n} g_N^{(j)}(x_1, x_2, \dots, x_n) O(x_1, \dots, x_n) dx_1 \dots dx_n, \end{aligned} \quad (21)$$

with the natural scaling $\eta = 1/N$, cannot be concluded from the fixed energy universality (20).

Given the historical importance of the Wigner surmise, it is somewhat surprising that gap universality with a fixed label did not receive much attention until very recently. The first results on the Wigner-Dyson-Gaudin-Mehta universality proved (20) in the sense of *average energy*, i.e. after taking average in the parameter E in a small interval of size $N^{-1+\varepsilon}$. Since $N^{-1+\varepsilon}$ is above the rigidity scale, average energy universality easily implies *average label gap universality*, i.e. the averaged version of (21) after averaging the label j in an interval of size N^ε .

Our more recent understanding shows that there is a profound difference between the weaker “*averaged*” results versus the stronger “*fixed*” ones. Obviously, “*fixed*” results are necessary for the precise statistics of individual points hence for fully characterizing the limiting process. At first sight, removing the local averaging may only seem a fine technical point; it merely requires to exclude the pathological case that a certain energy E (or a certain label j) might behave very differently than a typical one. Physicists have never worried about this situation since there is no apparent reason for such pathology (in fact Mehta’s original version of the conjecture did not specify the precise formulation of universality). Mathematically, however, it turned out surprisingly involved to exclude the worst case scenarios and we needed to develop a completely new approach. Finally, we point out that, unlike their averaged counterparts, the fixed energy and the fixed label results are not equivalent, in fact each required a separate proof.

5. Universality of local statistics: the main results

5.1. Wigner ensembles. Our main results hold for a larger class of ensembles than the standard Wigner matrices, which we will call *generalized Wigner matrices*.

Definition 5.1. ([36]) The real symmetric or complex Hermitian matrix ensemble H with centred and independent matrix elements $h_{ij} = \overline{h_{ji}}$, $i \leq j$, is called *generalized Wigner matrix* if the variances $\sigma_{ij}^2 = \mathbb{E}|h_{ij}|^2$ satisfy:

(A) For any j fixed

$$\sum_{i=1}^N \sigma_{ij}^2 = 1. \quad (22)$$

(B) There exist two positive constants, C_1 and C_2 , independent of N such that

$$\frac{C_1}{N} \leq \sigma_{ij}^2 \leq \frac{C_2}{N}. \quad (23)$$

For Hermitian ensembles, we additionally require that for each i, j the 2×2 covariance matrix is bounded by C/N in matrix sense, i.e.

$$\Sigma_{ij} := \begin{pmatrix} \mathbb{E}(\operatorname{Re} h_{ij})^2 & \mathbb{E}(\operatorname{Re} h_{ij})(\operatorname{Im} h_{ij}) \\ \mathbb{E}(\operatorname{Re} h_{ij})(\operatorname{Im} h_{ij}) & \mathbb{E}(\operatorname{Im} h_{ij})^2 \end{pmatrix} \geq \frac{C}{N}.$$

The following theorem settles the average energy version of the Wigner-Dyson-Gaudin-Mehta conjecture for generalized Wigner matrices. It is formulated under the weakest moment assumptions. The same result under somewhat more restrictive assumptions were already obtained in [32, 33]; see also [69] for the complex hermitian case and for a quite restricted class of real symmetric matrices. More details on the history can be found in [22].

Theorem 5.2 (Universality with averaged energy). *[28, Theorem 7.2] Suppose that $H = (h_{ij})$ is a complex Hermitian (respectively, real symmetric) generalized Wigner matrix. Suppose that for some constants $\varepsilon > 0$, $C > 0$,*

$$\mathbb{E} \left| \sqrt{N} h_{ij} \right|^{4+\varepsilon} \leq C. \quad (24)$$

Let $n \in \mathbb{N}$ and $O : \mathbb{R}^n \rightarrow \mathbb{R}$ be a test function (i.e. compactly supported and continuous). Fix $|E_0| < 2$ and $\xi > 0$, then with $b_N = N^{-1+\xi}$ we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int_{E_0 - b_N}^{E_0 + b_N} \frac{dE}{2b_N} \int_{\mathbb{R}^n} d\alpha_1 \cdots d\alpha_n O(\alpha_1, \dots, \alpha_n) \\ & \times \frac{1}{\varrho_{sc}(E)^n} \left(p_N^{(n)} - p_{G,N}^{(n)} \right) \left(E + \frac{\alpha_1}{N\varrho_{sc}(E)}, \dots, E + \frac{\alpha_n}{N\varrho_{sc}(E)} \right) = 0. \end{aligned} \quad (25)$$

Here ϱ_{sc} is the semicircle law defined in (12), $p_N^{(n)}$ is the n -point correlation function of the eigenvalue distribution of H (11), and $p_{G,N}^{(n)}$ is the n -point correlation function of an $N \times N$ GUE (respectively, GOE) matrix.

The additional rescaling in (25) with $\varrho_{sc}(E)$ is not essential, it just reflects the choice of variables under which the Gaussian correlation function is given exactly by the sine kernel (2) and not by some trivially rescaled version of it.

We remark that our method also provides an effective speed of convergence in (25). We also point out that the condition (23) can be relaxed, see Corollary 8.3 [27]. For example, the lower bound can be changed to $N^{-9/8+\varepsilon}$. Alternatively, under an additional symmetry condition on the law of the matrix elements, the upper bound can be relaxed to $N^{-8/9-\varepsilon}$.

For the next result, we introduce the notation $\llbracket A, B \rrbracket := \{A, A+1, \dots, B\}$ for any integers $A < B$. A relatively straightforward consequence of Theorem 5.2 is the average gap universality:

Corollary 5.3 (Gap universality with averaged label). *Let H be as in Theorem 5.2 and O be a test function of n variables. Fix small positive constants $\xi, \alpha > 0$. Then for any integer $j_0 \in \llbracket \alpha N, (1 - \alpha)N \rrbracket$ we have*

$$\lim_{N \rightarrow \infty} \frac{1}{2N^\xi} \sum_{|j-j_0| \leq N^\xi} [\mathbb{E} - \mathbb{E}^G] O(N(\lambda_j - \lambda_{j+1}), N(\lambda_j - \lambda_{j+2}), \dots, N(\lambda_j - \lambda_{j+n})) = 0. \quad (26)$$

Here λ_j 's are the ordered eigenvalues. \mathbb{E} and \mathbb{E}^G denote the expectation with respect to the Wigner ensemble H and the Gaussian (GOE or GUE) ensemble, respectively.

We remark that, similarly to the explicit formulas for the correlation functions (2), for Gaussian (GOE or GUE) ensembles there are explicit expressions for the gap distribution even without local averaging. They are given in terms of a Fredholm determinant of the corresponding kernel K , see [14, 16, 62].

Now we present our results for fixed energy:

Theorem 5.4 (Universality at fixed energy). *[11] Theorem 5.2 holds under the same conditions without averaging, i.e. for any E with $|E| < 2$ we have*

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^n} d\alpha_1 \cdots d\alpha_n O(\alpha_1, \dots, \alpha_n) \times \frac{1}{\varrho_{sc}(E)^n} \left(p_N^{(n)} - p_{G,N}^{(n)} \right) \left(E + \frac{\alpha_1}{N\varrho_{sc}(E)}, \dots, E + \frac{\alpha_n}{N\varrho_{sc}(E)} \right) = 0. \quad (27)$$

We remark that the fixed energy result (27) for the $\beta = 2$ (complex Hermitian) case was already known before, see [30, 70] for special cases and [34] for the general case. The $\beta = 2$ case is exceptional since the Harish-Chandra/Itzykson/Zuber identity allows one to compute correlation functions for Wigner matrices with a tiny Gaussian component. This method relies on an algebraic identity and cannot be generalized to other symmetry classes.

Finally, the gap universality with fixed label asserts that (26) holds without averaging.

Theorem 5.5 (Gap universality with fixed label). *[35, Theorem 2.2] Assuming subexponential decay of the matrix elements instead of (24), Corollary 5.3 holds without averaging:*

$$\lim_{N \rightarrow \infty} [\mathbb{E} - \mathbb{E}^G] O(N(\lambda_j - \lambda_{j+1}), N(\lambda_j - \lambda_{j+2}), \dots, N(\lambda_j - \lambda_{j+n})) = 0, \quad (28)$$

for any $j \in \llbracket \alpha N, (1 - \alpha)N \rrbracket$ with a fixed $\alpha > 0$.

More generally, for any $k, m \in \llbracket \alpha N, (1 - \alpha)N \rrbracket$ we have

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} O\left((N\varrho_k)(\lambda_k - \lambda_{k+1}), (N\varrho_k)(\lambda_k - \lambda_{k+2}), \dots, (N\varrho_k)(\lambda_k - \lambda_{k+n})\right) - \mathbb{E}^G O\left((N\varrho_m)(\lambda_m - \lambda_{m+1}), \dots, (N\varrho_m)(\lambda_m - \lambda_{m+n})\right) \right| = 0, \quad (29)$$

where the local density ϱ_k is defined by $\varrho_k := \varrho_{sc}(\gamma_k)$ with γ_k from (17).

The second part (29) of this theorem asserts that the gap distribution is not only independent of the specific Wigner ensemble, but it is also universal throughout the bulk spectrum. This is the counterpart of the statement that the appropriately rescaled correlation functions (27) have a limit that is independent of E , see (2).

Prior to our work, universality for a single gap was only achieved in the special case of the Gaussian unitary ensemble (GUE) in [71], which statement then easily implies the same results for complex Hermitian Wigner matrices satisfying the four moment matching condition.

5.2. Log-gases. In the case of invariant ensembles, it is well-known that for V satisfying certain mild conditions the sequence of one-point correlation functions, or densities, associated with $\mu = \mu_{\beta,V}^{(N)}$ from (7) has a limit as $N \rightarrow \infty$ and the limiting equilibrium density $\varrho_V(s)$ can be obtained as the unique minimizer of the functional

$$I(\nu) = \int_{\mathbb{R}} V(t)\nu(t)dt - \int_{\mathbb{R}} \int_{\mathbb{R}} \log |t - s| \nu(s)\nu(t)dtds.$$

We assume that $\varrho = \varrho_V$ is supported on a single compact interval, $[A, B]$ and $\varrho \in C^2(A, B)$. Moreover, we assume that V is *regular* in the sense that ϱ is strictly positive on (A, B) and vanishes as a square root at the endpoints, i.e.

$$\varrho(t) = s_A \sqrt{t - A} (1 + O(t - A)), \quad t \rightarrow A^+, \quad (30)$$

for some constant $s_A > 0$ and a similar condition holds at the upper edge.

It is known that these conditions are satisfied if, for example, V is strictly convex. In this case ϱ_V satisfies the equation

$$\frac{1}{2}V'(t) = \int_{\mathbb{R}} \frac{\varrho_V(s)ds}{t - s} \quad (31)$$

for any $t \in (A, B)$. For the Gaussian case, $V(x) = x^2/2$, the equilibrium density is given by the semicircle law, $\varrho_V = \varrho_{sc}$, see (12).

The following result was proven in Corollary 2.2 of [8] for convex real analytic potential V , it was generalized in Theorem 1.2 of [9] for the non-convex case and further generalized for arbitrary C^4 potential in Theorem 2.5 of [10].

Theorem 5.6 (Universality with averaged energy). *Assume $V \in C^4(\mathbb{R})$, regular and let $\beta > 0$. Consider the β -ensemble $\mu_V = \mu_{\beta,V}^{(N)}$ given in (7) with correlation functions $p_{V,N}^{(n)}$ defined analogously to (11). For the Gaussian case, $V(x) = x^2/2$, the correlation functions are denoted by $p_{G,N}^{(n)}$. Let $E_0 \in (A, B)$ lie in the interior of the support of ϱ and similarly let $E'_0 \in (-2, 2)$ be inside the support of ϱ_{sc} . Then*

for $b_N = N^{-1+\xi}$ with some $\xi > 0$ we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int d\alpha_1 \cdots d\alpha_n O(\alpha_1, \dots, \alpha_n) \\ & \times \left[\int_{E_0 - b_N}^{E_0 + b_N} \frac{dE}{2b_N} \frac{1}{\varrho(E)^n} p_{V,N}^{(n)} \left(E + \frac{\alpha_1}{N\varrho(E)}, \dots, E + \frac{\alpha_n}{N\varrho(E)} \right) \right. \\ & \quad \left. - \int_{E'_0 - b_N}^{E'_0 + b_N} \frac{dE'}{2b_N} \frac{1}{\varrho_{sc}(E')^n} p_{G,N}^{(n)} \left(E' + \frac{\alpha_1}{N\varrho_{sc}(E')}, \dots, E' + \frac{\alpha_n}{N\varrho_{sc}(E')} \right) \right] = 0, \end{aligned} \quad (32)$$

i.e. the correlation functions of $\mu_{\beta,V}^{(N)}$ averaged around E_0 asymptotically coincide with those of the Gaussian case. In particular, they are independent of E_0 .

Theorem 5.6 immediately implies gap universality with averaged label, exactly in the same way as Corollary 5.3 was deduced from Theorem 5.2; we refrain from stating it explicitly. The following two theorems show that these results hold without averaging.

Theorem 5.7 (Universality at fixed energy). *[11] Consider the setup of Theorem 5.6 and we additionally assume that $\beta \geq 1$. Then (32) holds without averaging, i.e. for any $E \in (A, B)$ and $E' \in (-2, 2)$ we have*

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int d\alpha_1 \cdots d\alpha_n O(\alpha_1, \dots, \alpha_n) \\ & \times \left[\frac{1}{\varrho(E)^n} p_{V,N}^{(n)} \left(E + \frac{\alpha_1}{N\varrho(E)}, \dots, E + \frac{\alpha_n}{N\varrho(E)} \right) \right. \\ & \quad \left. - \frac{1}{\varrho_{sc}(E')^n} p_{G,N}^{(n)} \left(E' + \frac{\alpha_1}{N\varrho_{sc}(E')}, \dots, E' + \frac{\alpha_n}{N\varrho_{sc}(E')} \right) \right] = 0. \end{aligned} \quad (33)$$

Prior to our work and with a different method, the same result was also proven in [63] for analytic potentials and for any $\beta > 0$ even if the support of ϱ has several intervals. An extension of the method to $V \in C^5$ is anticipated in [63].

To formulate the result for the gap universality with a fixed label, we define the quantiles $\gamma_{j,V}$ of the density ϱ_V by

$$\frac{j}{N} = \int_A^{\gamma_{j,V}} \varrho_V(x) dx, \quad (34)$$

similarly to (17). We set

$$\varrho_j^V := \varrho_V(\gamma_{j,V}), \quad \text{and} \quad \varrho_j := \varrho_{sc}(\gamma_j) \quad (35)$$

to be the limiting densities at the j -th quantiles. Let \mathbb{E}^{μ_V} and \mathbb{E}^G denote the expectation w.r.t. the measure μ_V and its Gaussian counterpart for $V(\lambda) = \frac{1}{2}\lambda^2$.

Theorem 5.8 (Gap universality with fixed label). *[35, Theorem 2.3] Consider the setup of Theorem 5.6 and we also assume $\beta \geq 1$. Set some $\alpha > 0$, then*

$$\lim_{N \rightarrow \infty} \left| \mathbb{E}^{\mu_V} O\left((N \varrho_k^V)(\lambda_k - \lambda_{k+1}), (N \varrho_k^V)(\lambda_k - \lambda_{k+2}), \dots, (N \varrho_k^V)(\lambda_k - \lambda_{k+n})\right) - \mathbb{E}^{\mu_G} O\left((N \varrho_m)(\lambda_m - \lambda_{m+1}), \dots, (N \varrho_m)(\lambda_m - \lambda_{m+n})\right) \right| = 0 \quad (36)$$

for any $k, m \in [\alpha N, (1 - \alpha)N]$. In particular, the distribution of the rescaled gaps w.r.t. μ_V does not depend on the index k in the bulk.

We point out that Theorem 5.6 holds for any $\beta > 0$, but Theorems 5.7 and 5.8 require $\beta \geq 1$. This is only a technical restriction related to a certain condition in the De Giorgi-Nash-Moser regularity theory that is the backbone of our proof. Indeed, a year after our work was completed, an alternative proof of (36) was given for any $\beta > 0$ but with a higher regularity assumption on V and with an additional hypothesis that can be effectively checked only for convex V , see [4].

5.3. Universalities at the edge. We stated our results for the bulk of the spectrum. Similar results hold at the edge; in this case the “averaged” results are meaningless. For completeness, we give the universality results for both ensembles.

Theorem 5.9 (Universality at the edge for Wigner matrices). *[10] Let H be a generalized Wigner ensemble with subexponentially decaying matrix elements. Fix $n \in \mathbb{N}$, $\kappa < 1/4$ and a test function O of n variables. Then for any $\Lambda \subset [1, N^\kappa]$ with $|\Lambda| = n$, we have*

$$\left| [\mathbb{E} - \mathbb{E}^G] O\left(\left(N^{2/3} j^{1/3} (\lambda_j - \gamma_j)\right)_{j \in \Lambda}\right) \right| \leq N^{-\chi},$$

with some $\chi > 0$, where \mathbb{E}^G is expectation w.r.t. the standard GOE or GUE ensemble depending on the symmetry class of H and γ_j ’s are semicircle quantiles.

Edge universality for Wigner matrices was first proved in [67] assuming symmetry of the distribution of the matrix elements and finiteness of all their moments. The symmetry condition was completely eliminated [37] and the optimal moment condition was obtained in [48]. All these works heavily rely on the fact that the variances of the matrix elements are identical. The main point of Theorem 5.9 is to consider generalized Wigner matrices, i.e., matrices with non-constant variances. In fact, it was shown in [37] that the edge statistics for any generalized Wigner matrix are universal in the sense that they coincide with those of a generalized Gaussian Wigner matrix with the same variances, but it was not shown that the statistics are independent of the variances themselves. Theorem 5.9 provides this missing step and thus it proves the edge universality in the broadest sense.

Theorem 5.10 (Universality at the edge for log-gases). *[10] Let $\beta \geq 1$ and V (resp. \tilde{V}) be in $C^4(\mathbb{R})$, regular such that the equilibrium density ϱ_V (resp. $\varrho_{\tilde{V}}$) is supported on a single interval $[A, B]$ (resp. $[\tilde{A}, \tilde{B}]$). Without loss of generality we assume that for both densities (30) holds with $A = 0$ and with the same constant s_A . Fix $n \in \mathbb{N}$, $\kappa < 2/5$. Then for any $\Lambda \subset \llbracket 1, N^\kappa \rrbracket$ with $|\Lambda| = n$, we have*

$$\left| (\mathbb{E}^{\mu_V} - \mathbb{E}^{\mu_{\tilde{V}}}) O \left(\left(N^{2/3} j^{1/3} (\lambda_j - \gamma_j) \right)_{j \in \Lambda} \right) \right| \leq N^{-\chi} \quad (37)$$

with some $\chi > 0$. Here γ_j are the quantiles w.r.t. the density ϱ_V (34).

The first results on edge universality for invariant ensembles concerned the classical values of $\beta = 1, 2, 4$. The case $\beta = 2$ and real analytic V was solved in [17, 15]. The $\beta = 1, 4$ cases are considerably harder than $\beta = 2$. For $\beta = 1, 4$ universality was first proved for polynomial potentials in [15], then for the real analytic case for $\beta = 1$ in [59, 61], which also give an alternative proof for $\beta = 2$. Finally, independently of our work with a completely different method, edge universality for any $\beta > 0$ and convex polynomial V was recently proved in [46].

6. Outline of the proof strategy

6.1. “Averaged” results: Dyson Brownian motion. The proof of Theorem 5.2 follows a three-step strategy that was first introduced in [30] and further developed in [32].

Step 1. Local semicircle law and rigidity of eigenvalues. The main tool is the resolvent of H at a spectral parameter $z = E + i\eta$ with $\eta \gg 1/N$;

$$m_N(z) := \frac{1}{N} \text{Tr} \frac{1}{H - z} = \frac{1}{N} \sum_j \frac{1}{\lambda_j - z},$$

which is of the form of (16) with $O(x) = (x - i)^{-1}$. Using the Schur decomposition formula we may write

$$m_N(z) = \frac{1}{N} \sum_{j=1}^N \frac{1}{h_{jj} - z - \sum_{a,b \neq j} h_{ja} G_{ab}^{(j)}(z) h_{bj}},$$

where $G^{(j)}(z) = (H^{(j)} - z)^{-1}$ is the resolvent of the $(N-1) \times (N-1)$ minor $H^{(j)}$ of H after removing the j -th row and column. Since $G_{ab}^{(j)}(z)$ and $h_{ja} h_{bj}$ are independent, we may use concentration results to replace the double sum in the denominator by its expectation over the matrix elements in the j -th row and column. Neglecting the fluctuation, we recover $m_N^{(j)}(z)$, the normalized trace of the resolvent of $H^{(j)}$. Since $m_N^{(j)}(z)$ and $m_N(z)$ are close, we obtain the following *self-consistent equation*

$$m_N(z) = -\frac{1}{z + m_N(z)} + \text{error}. \quad (38)$$

If the error is neglected, then the solution of the resulting quadratic equation is exactly the Stieltjes transform

$$m_{sc}(z) := \int_{\mathbb{R}} \frac{1}{x-z} \varrho_{sc}(x) dx$$

of the Wigner semicircle law $\varrho_{sc}(x)$. This allows us to conclude that $m_N(z)$ is close to m_{sc} , and a careful analysis yields

$$|m_N(z) - m_{sc}(z)| \lesssim \frac{1}{N\eta}. \quad (39)$$

This is the local semicircle law in resolvent form, from which the limit of (16) and the rigidity property (18) can be concluded.

Step 2. Universality for Gaussian divisible ensembles: The Gaussian divisible ensembles are matrices of the form

$$H_t = e^{-t/2} H + \sqrt{1 - e^{-t}} U,$$

where H is a Wigner matrix and U is an independent GUE/GOE matrix. The parametrization of H_t reflects that, in the sense of distribution, it is most conveniently obtained by an Ornstein-Uhlenbeck process:

$$dH_t = \frac{1}{\sqrt{N}} dB_t - \frac{1}{2} H_t dt, \quad (40)$$

where B_t is a matrix-valued Brownian motion of the appropriate symmetry class. Dyson observed [20] that the corresponding process λ_t of the eigenvalues of H_t remarkably satisfies a system of stochastic differential equations (SDE), called the *Dyson Brownian Motion (DBM)*:

$$d\lambda_j = \frac{1}{\sqrt{N}} dB_j + \left[-\frac{1}{2} \lambda_j + \frac{1}{N} \sum_{k \neq j} \frac{1}{\lambda_k - \lambda_j} \right] dt, \quad (41)$$

written for $\beta = 2$, where B_j 's are independent standard real Brownian motions. The key idea is to study the relaxation of the flow (41) to its equilibrium measure which is the distribution of the GUE eigenvalues. It turns out that, tested against observables involving only *differences of eigenvalues*, the convergence is extremely fast. Combined with the rigidity bound that guarantees a strong apriori control on the initial state, we obtain that the gap statistics are already in local equilibrium (hence universal) after a very short time $t = N^{-1+\varepsilon}$, see [32, 33].

This method substantially improves Johansson's result [44] which showed universality only with a substantial Gaussian component (essentially for $t > 0$ independent of N) and only for the $\beta = 2$ symmetry class. In fact, the first restriction can be relaxed by using our optimal rigidity bound [30, 34], but the second one cannot be removed since the proof relies on the Harish-Chandra/Itzykson/Zuber formula. The analysis of the DBM is much more robust, in particular it applies to any symmetry class. However, it yields only an averaged result (26) (from which

(25) can be deduced), while [30, 34] gives the fixed energy results (27) but only for $\beta = 2$.

Step 3. Approximation by Gaussian divisible ensembles: It is a simple density argument in the space of matrix ensembles which shows that for any probability distribution of the matrix elements there exists a Gaussian divisible distribution with a small Gaussian component, as in Step 2, such that the two associated Wigner ensembles have asymptotically identical local eigenvalue statistics. The first implementation of this approximation scheme was via a reverse heat flow argument [30]; it was later replaced by the *Green function comparison theorem* [36] motivated by the four moment matching condition of [69]. This comparison argument is very robust: it works even without averaging and for arbitrary observables not only for those of difference type.

The proof of Theorem 5.6 follows a somewhat similar path but with essential differences. Rigidity estimates still hold on the smallest scale, but their derivation cannot use resolvents since there is no matrix behind a general log-gas. Instead of (38) we use the loop equation from [45] or [62], but extended to smooth potentials. There is no analogue of the Gaussian divisible ensemble for log-gases, but an enhanced version of the DBM underlying the invariant measure μ_V can still be analyzed.

In summary, the DBM plays the fundamental role behind the “averaged” universality result for both models.

6.2. “Fixed” results: Hölder regularity and homogenization.

For definiteness, we will present some ideas to prove Theorem 5.8, the proof of Theorems 5.4, 5.5, 5.7 and the results at the edge require additional steps.

Step 1. Comparison of local Gibbs measures. The basic mechanism for universality is that the microscopic structure of the measure μ_V defined in (6) is insensitive of the potential V , it is essentially determined by the Vandermonde determinant, i.e. the log-interaction in (7). In the first step we localize the problem by freezing (conditioning on) all particles at a distance $1 \ll K \ll N$ away from the fixed index j of the gap $\lambda_j - \lambda_{j+1}$ we want to study. Thus the corresponding local Gibbs measure is defined on an interval $I = [j - K, j + K]$ and it still retains the Vandermonde structure. On this mesoscopic scale the potential is locally constant, hence its effect is trivial, so the key question is to show that $\lambda_j - \lambda_{j+1}$ is largely insensitive to the boundary effects we just introduced by localization. This is a question about the long range correlation structure of the Gibbs measure.

The main difficulty is that the log-gas is a strongly correlated system in contrast to the customary setup in statistical physics where correlations often decay very fast. In fact, the covariance between two points decays only logarithmically

$$\frac{\langle \lambda_i; \lambda_j \rangle}{\sqrt{\langle \lambda_i; \lambda_i \rangle \langle \lambda_j; \lambda_j \rangle}} \sim \frac{1}{\log |i - j|}, \quad 1 \ll |i - j| \ll N. \quad (42)$$

One key observation is that the correlation decay between a *gap* $\lambda_i - \lambda_{i+1}$ and a *point* λ_j is faster, it is $|i - j|^{-1}$, practically the discrete derivative of (42).

Step 2. Random walk representation of the covariance. In a more general setup, consider a Gibbs measure $\omega(d\mathbf{x}) = e^{-\beta\mathcal{H}(\mathbf{x})}d\mathbf{x}$ on finitely many points labelled by I and with a strictly convex Hamiltonian, $\mathcal{H}''(\mathbf{x}) \geq c > 0$. Then the covariance w.r.t. ω can be expressed as

$$\langle F(\mathbf{x}); G(\mathbf{x}) \rangle_\omega = \frac{1}{2} \int_0^\infty ds \int d\omega(\mathbf{x}) \mathbb{E}_\mathbf{x} [\nabla G(\mathbf{x}(s)) \cdot \mathcal{U}(s, \mathbf{x}(\cdot)) \nabla F(\mathbf{x})], \quad (43)$$

see [43, 55]. Here $\mathbb{E}_\mathbf{x}$ is the expectation for the (random) paths $\mathbf{x}(\cdot)$ starting from $\mathbf{x}(0) = \mathbf{x}$ and solving the canonical SDE for the measure ω :

$$d\mathbf{x}(s) = d\mathbf{B}(s) - \beta \nabla \mathcal{H}(\mathbf{x}(s)) ds, \quad (44)$$

and $\mathcal{U}(s) = \mathcal{U}(s, \mathbf{x}(\cdot))$ is the fundamental solution to the linear system of equations

$$\partial_s \mathcal{U}(s) = -\mathcal{U}(s) \mathcal{A}(s), \quad \mathcal{A}(s) := \beta \mathcal{H}''(\mathbf{x}(s)) \quad (45)$$

with $\mathcal{U}(0) = \text{Id}$. Notice that the coefficient matrix $\mathcal{A}(s)$, and thus the fundamental solution, depend on the random path. The SDE (44) is the generalization of the DBM, (41). Formula (43) turns the problem of computing the covariance $\langle F; G \rangle$ into a time-dependent question to understand the fundamental solution \mathcal{U} of the parabolic equation (45).

In particular, if G is a function of a single gap, $G(\mathbf{x}) = O(x_j - x_{j+1})$ with some fixed j , and F represents the boundary effects, then (43) becomes

$$\frac{1}{2} \int_0^\infty ds \int d\omega(\mathbf{x}) \sum_{i \in I} \mathbb{E}_\mathbf{x} [O'(x_j - x_{j+1}) (\mathcal{U}_{i,j}(s) - \mathcal{U}_{i,j+1}(s)) \partial_i F(\mathbf{x})]. \quad (46)$$

The key technical step is to show that for a typical path $\mathbf{x}(\cdot)$ the solution $\mathcal{U}(s)$ is Hölder-regular in a sense that $\mathcal{U}_{i,j}(s) - \mathcal{U}_{i,j+1}(s)$ is small if j is away from the boundary of I and s is not too small.

Step 3. Hölder-regularity of the solution to (45). For any fixed realization of the path $\mathbf{x}(\cdot)$, we will view the equation (45) as a finite dimensional version of a parabolic equation. The coefficient matrix, the Hessian of the local Gibbs measure, is computed explicitly. It can be written as $\mathcal{A} = \mathcal{B} + \mathcal{W}$, where $\mathcal{W} \geq 0$ is diagonal, \mathcal{B} is a symmetric matrix with quadratic form

$$\langle \mathbf{u}, \mathcal{B}(s) \mathbf{u} \rangle = \frac{1}{2} \sum_{i,j \in I} B_{ij}(s) (u_i - u_j)^2, \quad B_{ij}(s) := \frac{\beta}{(x_i(s) - x_j(s))^2}.$$

After rescaling the problem so that the gap is of order one, for a typical path and large $i - j$ we have

$$B_{ij}(s) \sim \frac{1}{(i - j)^2} \quad (47)$$

by rigidity. We also have a lower bound for any $i \neq j$

$$B_{ij}(s) \gtrsim \frac{1}{(i - j)^2}, \quad (48)$$

at least with a very high probability. If a matching upper bound were true for any $i \neq j$, then (45) would be the discrete analogue of the general equation

$$\partial_t u(t, x) = \int K(t, x, y)[u(t, y) - u(t, x)]dy, \quad t > 0, \quad x, y \in \mathbb{R}^d \quad (49)$$

considered by Caffarelli-Chan-Vasseur in [12], where the kernel K is symmetric and has a specific short distance singularity

$$C_1|x - y|^{-d-s} \leq K(t, x, y) \leq C_2|x - y|^{-d-s} \quad (50)$$

for some $s \in (0, 2)$ and positive constants C_1, C_2 . Roughly speaking, the integral operator K corresponds to the behavior of the operator $|p|^s$, where $p = -i\nabla$. The main result of [12] asserts that for any $t_0 > 0$, the solution $u(t, x)$ is ε -Hölder continuous, $u \in C^\varepsilon((t_0, \infty), \mathbb{R}^d)$, for some positive exponent ε that depends only on t_0, C_1, C_2 . This is a version of the celebrated De Giorgi-Nash-Moser regularity result for a non-local operator.

Our equation (45) is of this type with $d = s = 1$, but it is discrete and in a finite interval I with a potential term. The key difference, however, is that the coefficient $B_{ij}(t)$ can be singular in the sense that $B_{ij}(t)|i - j|^2$ is not uniformly bounded when i, j are close to each other. Thus the analogue of the uniform upper bound (50) does not even hold for a fixed t . We first need to regularize the singularity of B_{ij} on a very tiny scale. Even after that we can control the regularized B_{ij}^{reg} only in a certain average sense:

$$\sup_{0 \leq s \leq \sigma} \sup_{0 \leq M \leq K} \frac{1}{1+s} \int_0^s \frac{1}{M} \sum_{i \in I: |i-Z| \leq M} B_{i,i+1}^{\text{reg}}(s) ds \leq CK^\rho \quad (51)$$

with high probability, for some small exponent ρ and for any fixed Z away from the edges of I . This estimate essentially says that the space-time maximal function of $B_{i,i+1}^{\text{reg}}(t)$ at a fixed space-time point $(Z, 0)$ is bounded by K^ρ . Our main generalization of the result in [12] is to show that the weak upper bound (51) at a few space-time points together with (47) and (48) (holding up to a factor K^ξ) are sufficient for proving a discrete version of the Hölder continuity at the point $(Z, 0)$. More precisely, there exists an $\varepsilon > 0$ such that for any fixed $1 \ll \sigma \ll K$ the solution to (45) satisfies

$$\sup_{|j-Z|+|j'-Z| \leq \sigma^{1-\alpha}} |\mathcal{U}_{i,j}(\sigma) - \mathcal{U}_{i,j'}(\sigma)| \leq CK^\xi \sigma^{-1-\varepsilon\alpha} \quad (52)$$

with any $\alpha \in [0, 1/3]$ if we can guarantee that ρ and ξ are sufficiently small. The exponent ε plays the role of the Hölder regularity exponent. Notice that $\mathcal{U}_{i,j}(\sigma)$ decays as σ^{-1} , hence (52) provides an additional decay for the discrete derivative. In particular, this guarantees that the ds integration in (46) is finite. With several further technical steps, this proves Theorem 5.8.

Step 4. Homogenization. The proofs of Theorems 5.5, 5.7 require an additional information about the fundamental solution of (45). Since in the $|i - j| \gg 1$ regime

we have $B_{ij}(s) \sim |i-j|^{-2}$, it is reasonable to expect that the large time and large scale behavior of \mathcal{U} is given by the

$$\mathcal{U}_{ij}(t) \approx \left(e^{-t|p|} \right)_{ij} = \frac{t}{t^2 + (i-j)^2}, \quad |i-j| \gg 1, \quad t \gg 1, \quad (53)$$

where we computed the heat kernel of $|p| = \sqrt{-\Delta}$ explicitly. This result, combined with a coupling argument, yields that

$$\lambda_i(t) - \tilde{\lambda}_i(t) = \left(e^{-t|p|} \boldsymbol{\lambda}(0) \right)_i - \left(e^{-t|p|} \tilde{\boldsymbol{\lambda}}(0) \right)_i + \text{error}, \quad (54)$$

where $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ are two solutions of the SDE (44) with the same Brownian motion $\mathbf{B}(s)$ but with two different initial conditions. In the applications, $\boldsymbol{\lambda}(0)$ will be GUE/GOE eigenvalues and $\tilde{\boldsymbol{\lambda}}(0)$ will be the eigenvalues of a general Wigner matrix. Formula (54) allows us to express a single Wigner eigenvalue $\lambda_i(t)$ in terms of the corresponding Gaussian eigenvalue $\tilde{\lambda}_i(t)$ and in terms of averaged quantities involving many eigenvalues. Since averaged quantities can be computed much easier and Gaussian computations can be performed by explicit formulas, we obtain nontrivial information about $\lambda_i(t)$. Finally, approximation ideas similar to Step 3. in Section 6.1 can relate general Wigner eigenvalues to Wigner eigenvalues with some Gaussian component such as $\lambda_i(t)$. In particular, these ideas can prove the logarithmic correlation decay (42) for any Wigner matrix.

In summary, the detailed analysis of the parabolic equation (45) with singular coefficients given by the Dyson Brownian motion play the crucial role behind all “fixed” universality results for both Wigner matrices and log-gases.

Acknowledgement. Most results in this paper were obtained in collaboration with Horng-Tzer Yau, Benjamin Schlein, Jun Yin, Antti Knowles and Paul Bourgade and in some work, also with Jose Ramirez and Sandrine Péche. This article reports the joint progress with these authors.

References

- [1] Aizenman, M., and Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation, *Commun. Math. Phys.* **157**, 245–278 (1993)
- [2] Ajanki, O., Erdős, L., Krüger, T.: Universality for Wigner random matrices with non-stochastic variance matrix. Preprint.
- [3] Anderson, P.: Absences of diffusion in certain random lattices, *Phys. Rev.* **109**, 1492–1505 (1958)
- [4] Bekerman, F., Figalli, A., Guionnet, A.: Transport maps for β -matrix models and universality. Preprint. arXiv:1311.2315
- [5] Bleher, P., Its, A.: Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. *Ann. of Math.* **150** (1999), 185–266.

- [6] Boutet de Monvel, A., Khorunzhy, A.: Asymptotic distribution of smoothed eigenvalue density. I. Gaussian random matrices. *Random Oper. and Stoch. Equ.* **7**, 1–22 (1999)
- [7] Boutet de Monvel, A., Khorunzhy, A.: Asymptotic distribution of smoothed eigenvalue density. II. Wigner random matrices. *Random Oper. and Stoch. Equ.* **7**, 149–168 (1999)
- [8] Bourgade, P., Erdős, L., Yau, H.-T.: Universality of General β -Ensembles, arXiv:1104.2272
- [9] Bourgade, P., Erdős, L., Yau, H.-T.: Bulk Universality of General β -Ensembles with Non-convex Potential, *J. Math. Phys.* **53**, 095221 (2012)
- [10] Bourgade, P., Erdős, L., Yau, H.-T.: Edge universality of beta ensembles, Preprint, arXiv:1306.5728
- [11] Bourgade, P., Erdős, L., Yau, H.-T., Yin, J.: Fixed energy universality for generalized Wigner matrices. Preprint. arXiv:1407.5606
- [12] Caffarelli, L., Chan, C.H., Vasseur, A.: Regularity theory for parabolic nonlinear integral operators, *J. Amer. Math. Soc.* **24**, no. 3, 849–889 (2011)
- [13] Deuschel, J.-D., Giacomin, G., Ioffe, D.: Large deviations and concentration properties for $\nabla\varphi$ interface models. *Probab. Theor. Relat. Fields.* **117**, 49–111 (2000)
- [14] Deift, P.: Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. *Courant Lecture Notes in Mathematics* **3**, American Mathematical Society, Providence, RI, 1999
- [15] Deift, P., Gioev, D.: Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices. *Comm. Pure Appl. Math.* **60**, no. 6, 867–910 (2007)
- [16] Deift, P., Gioev, D.: Random Matrix Theory: Invariant Ensembles and Universality. *Courant Lecture Notes in Mathematics* **18**, American Mathematical Society, Providence, RI, 2009
- [17] Deift, P., Kriecherbauer, T., McLaughlin, K.T-R, Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52**, 1335–1425 (1999)
- [18] Deift, P., Kriecherbauer, T., McLaughlin, K.T-R, Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52**, 1491–1552 (1999)
- [19] Dyson, F.J.: Statistical theory of energy levels of complex systems, I, II, and III. *J. Math. Phys.* **3**, 140–156, 157–165, 166–175 (1962)
- [20] Dyson, F.J.: A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3**, 1191–1198 (1962)
- [21] Dyson, F.J.: Correlations between eigenvalues of a random matrix. *Commun. Math. Phys.* **19**, 235–250 (1970)
- [22] Erdős, L.: *Universality for random matrices and log-gases. In: Current Developments in Mathematics 2012*, Ed. D. Jerison, M. Kisin, T. Mrowka. R. Stanley, H.-T. Yau, S.-T. Yau, International Press, 59–132 (2013)

- [23] Erdős, L., Knowles, A.: Quantum Diffusion and Eigenfunction Delocalization in a Random Band Matrix Model. *Commun. Math. Phys.* **303** no. 2, 509–554 (2011)
- [24] Erdős, L., A. Knowles, A.: Quantum Diffusion and Delocalization for Band Matrices with General Distribution. *Annales Inst. H. Poincaré*, **12** (7), 1227–1319 (2011)
- [25] Erdős, L., Knowles, A.: *The Altshuler-Shklovskii formulas for random band matrices I: the unimodular case*. To appear in Commun. Math. Phys. [arXiv:1309.5106](#)
- [26] Erdős, L., Knowles, A.: *The Altshuler-Shklovskii formulas for random band matrices II: the general case*. To appear in Annals Henri Poincaré. [arXiv:1309.5107](#)
- [27] Erdős, L., A. Knowles, A., Yau, H.-T., J. Yin.: The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** no. 59, 1–58. (2013)
- [28] Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Spectral Statistics of Erdős-Rényi Graphs II: Eigenvalue Spacing and the Extreme Eigenvalues. *Comm. Math. Phys.* **314** no. 3. 587–640 (2012)
- [29] Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Delocalization and Diffusion Profile for Random Band Matrices. Preprint: [arXiv:1205.5669](#)
- [30] Erdős, L., Péché, G., Ramírez, J., Schlein, B., and Yau, H.-T., Bulk universality for Wigner matrices. *Commun. Pure Appl. Math.* **63**, No. 7, 895–925 (2010)
- [31] Erdős, L., Schlein, B., Yau, H.-T.: Local semicircle law and complete delocalization for Wigner random matrices. *Commun. Math. Phys.* **287**, 641–655 (2009)
- [32] Erdős, L., Schlein, B., Yau, H.-T.: Universality of random matrices and local relaxation flow. *Invent. Math.* **185** (2011), no.1, 75–119.
- [33] Erdős, L., Schlein, B., Yau, H.-T., Yin, J.: The local relaxation flow approach to universality of the local statistics for random matrices. *Annales Inst. H. Poincaré (B), Probability and Statistics* **48**, no. 1, 1–46 (2012)
- [34] Erdős, L., Yau, H.-T.: A comment on the Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices. *Electron. J. Probab.* **17**, no 28. 1–5 (2012)
- [35] Erdős, L., Yau, H.-T.: Gap universality of generalized Wigner and β -ensembles. Preprint [arXiv:1211.3786](#)
- [36] Erdős, L., Yau, H.-T., Yin, J.: Bulk universality for generalized Wigner matrices. *Prob. Theor. Rel. Fields*, **154**, no. 1-2., 341–407 (2012)
- [37] Erdős, L., Yau, H.-T., Yin, J.: Rigidity of Eigenvalues of Generalized Wigner Matrices. *Adv. Math.* **229**, no. 3, 1435–1515 (2012)
- [38] Fokas, A. S., Its, A. R., Kitaev, A. V.: The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.* **147**, 395–430 (1992)
- [39] Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy, *Commun. Math. Phys.* **88**, 151–184 (1983)
- [40] Fyodorov, Y.V. and Mirlin, A.D.: Scaling properties of localization in random band matrices: a σ -model approach. *Phys. Rev. Lett.* **67**, 2405–2409 (1991)
- [41] Gaudin, M.: Sur la loi limit de l’espacement des valeurs propres d’une matrice aléatoire. *Nucl. Phys.* **25**, 447–458.
- [42] Gustavsson, J. : Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. H. Poincaré Probab. Statist.* **41**, no. 2, 151–178 (2005)

- [43] Helffer, B., Sjöstrand, J.: On the correlation for Kac-like models in the convex case. *J. Statis. Phys.* **74**, no.1-2, 349–409 (1994)
- [44] Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* **215**, no.3. 683–705 (2001)
- [45] Johansson, K.: On the fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91**, 151–204 (1998)
- [46] Krishnapur, M., Rider, B., Virág, B.: Universality of the Stochastic Airy Operator. Preprint arXiv:1306.4832
- [47] Kriecherbauer, T., Shcherbina, M.: Fluctuations of eigenvalues of matrix models and their applications. Preprint arXiv:1003.6121
- [48] Lee, J. O., Yin, J.: A Necessary and Sufficient Condition for Edge Universality of Wigner matrices. Preprint. arXiv:1206.2251
- [49] Lubinsky, D.S.: A New Approach to Universality Limits Involving Orthogonal Polynomials, *Ann. Math.*, **170**, 915–939 (2009)
- [50] Mehta, M.L.: *Random Matrices*. Third Edition, Academic Press, New York, 1991.
- [51] Mehta, M.L.: A note on correlations between eigenvalues of a random matrix. *Commun. Math. Phys.* **20** no.3. 245–250 (1971)
- [52] Mehta, M.L., Gaudin, M.: On the density of eigenvalues of a random matrix. *Nuclear Phys.* **18**, 420–427 (1960).
- [53] Minami, N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Commun. Math. Phys.* **177**, 709–725 (1996)
- [54] Montgomery, H.L.: The pair correlation of zeros of the zeta function. Analytic number theory, Proc. of Sympos. in Pure Math. **24**, Amer. Math. Soc. Providence, R.I., 181–193 (1973).
- [55] Naddaf, A., Spencer, T.: On homogenization and scaling limit of some gradient perturbations of a massless free field, *Commun. Math. Phys.* **183**, no.1., 55–84 (1997)
- [56] O’Rourke, S.: Gaussian Fluctuations of Eigenvalues in Wigner Random Matrices. *J. Statis. Phys.* **138**, no. 6, 1045–1066 (2010)
- [57] Pastur, L., Shcherbina, M.: Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Stat. Phys.* **86**, 109–147 (1997)
- [58] Pastur, L., Shcherbina M.: Bulk universality and related properties of Hermitian matrix models. *J. Stat. Phys.* **130**, no.2., 205–250 (2008)
- [59] Pastur, L., Shcherbina M.: On the edge universality of the local eigenvalue statistics of matrix model. *Mat. Fiz. Anal. Geom.* **10**, no. 3, 335–365 (2003)
- [60] Schenker, J.: Eigenvector localization for random band matrices with power law band width. *Commun. Math. Phys.* **290**, 1065–1097 (2009)
- [61] Shcherbina, M.: Edge universality for orthogonal ensembles of random matrices. *J. Stat. Phys.* **136**, no. 1, 35–50 (2009)
- [62] Shcherbina, M.: Orthogonal and symplectic matrix models: universality and other properties. *Comm. Math. Phys.* **307**, no.3., 761–790 (2011)
- [63] Shcherbina, M.: Change of variables as a method to study general β -models: bulk universality. Preprint. arXiv:1310.7835

- [64] Shcherbina, T.: On the second mixed moment of the characteristic polynomials of the 1D band matrices. Preprint. arXiv:1209.3385
- [65] Shcherbina, T.: Universality of the local regime for the block band matrices with a finite number of blocks. Preprint. arXiv:1309.2120
- [66] Sodin, S.: An estimate for the average spectral measure of random band matrices, *J. Stat. Phys.* **144**, no. 1, 46–59 (2011)
- [67] Soshnikov, A.: Universality at the edge of the spectrum in Wigner random matrices, *Comm. Math. Phys.*, **207**, no.3, 697–733 (1999)
- [68] Spencer, T.: Random banded and sparse matrices (Chapter 23), to appear in “Oxford Handbook of Random Matrix Theory” edited by G. Akemann, J. Baik, and P. Di Francesco
- [69] Tao, T. and Vu, V.: Random matrices: Universality of the local eigenvalue statistics. *Acta Math.*, **206**, no. 1, 127–204 (2011)
- [70] Tao, T. and Vu, V.: The Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices. *Electron. J. Probab.* **16**, no 77, 2104–2121 (2011)
- [71] Tao, T.: The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. Preprint. arXiv:1203.1605
- [72] Tracy, C., Widom, H.: Level spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159**, 151–174, 1994.
- [73] Tracy, C., Widom, H.: On orthogonal and symplectic matrix ensembles, *Comm. Math. Phys.* **177** (1996), no. 3, 727–754.
- [74] Valkó, B.; Virág, B.: Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177**, no. 3, 463–508 (2009)
- [75] Wigner, E.: Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* **62**, 548–564 (1955)

IST Austria, Am Campus 1, A-3400 Klosterneuburg, Austria
 E-mail: lerdos@ist.ac.at